

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-024192
 (43)Date of publication of application : 25.01.2002

(51)Int.Cl.

G06F 15/177

G06F 9/46

G06F 13/00

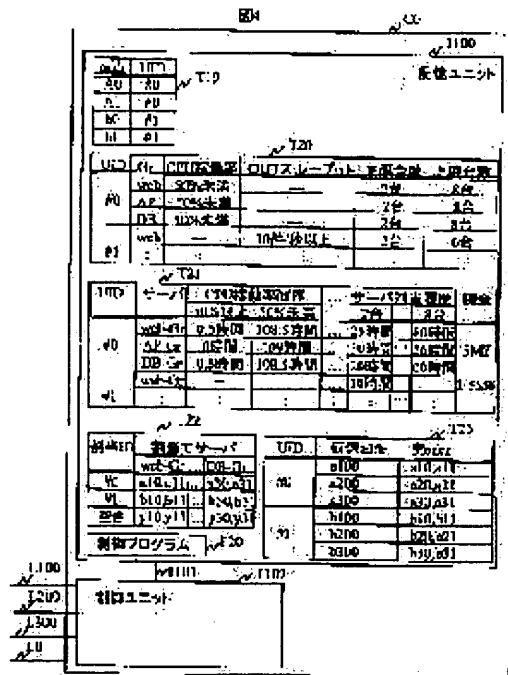
(21)Application number : 2000-211980

(71)Applicant : HITACHI LTD

(22)Date of filing : 07.07.2000

(72)Inventor : TAMAOKI YOSHIKO
 SHONAI TORU
 SAGAWA NOBUTOSHI
 KAWABE TAKASHI

(54) DEVICE AND METHOD FOR DIVIDING COMPUTER RESOURCES



(57)Abstract:

PROBLEM TO BE SOLVED: To allocate the independent calculation resources to each user corporation via a data center and to change automatically and in real time the allocation of calculation resources according to the load.

SOLUTION: A control program P20 prepares computer allocation control tables (T22, T23) for each user corporation identifier (#0, etc.), and sets these tables on a load distribution device d100. A means (T19) identifies the user corporation identifiers among the IP addresses (A0, etc.), contained in a user request packet. The device d100 identifies a user corporation from the user request packet and transfers the corporation to a due computer group. Then the program P20 compares the service level contract (T20) of each user with the operating state monitor result of a computer (CPU usage rate history contained in T21) and increases the allocated computers when the contract condition is not satisfied. In other words, the tables T22 and T23 are changed and set on the device d100. Furthermore, the program P20 changes the accounting information (server allocation history contained in T21 and accounting) based on the changed and set tables T22 and T23.

LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2002-24192

(P2002-24192A)

(43)公開日 平成14年1月25日 (2002.1.25)

(51) Int.Cl.
 G 0 6 F 15/177
 9/46
 13/00

識別記号
 6 7 4
 3 6 0
 3 5 7

F I
 G 0 6 F 15/177
 9/46
 13/00

テ-マコト(参考)
 6 7 4 A 5 B 0 4 5
 3 6 0 C 5 B 0 8 9
 3 5 7 Z 5 B 0 9 8

審査請求 未請求 請求項の数28 O L (全23頁)

(21)出願番号

特願2000-211980(P2000-211980)

(22)出願日

平成12年7月7日 (2000.7.7)

(71)出願人 000005108

株式会社日立製作所
東京都千代田区神田駿河台四丁目6番地

(72)発明者 玉置 由子

東京都国分寺市東恋ヶ窪一丁目280番地
株式会社日立製作所中央研究所内

(72)発明者 庄内 亨

東京都国分寺市東恋ヶ窪一丁目280番地
株式会社日立製作所中央研究所内

(74)代理人 100075096

弁理士 作田 康夫

最終頁に続く

(54)【発明の名称】 計算機資源分割装置および資源分割方法

(57)【要約】

【課題】データセンタにおいて、ユーザ企業ごとに夫々独立な計算資源を割り当てる、負荷に応じてその割当てを自動的にリアルタイムに変更する

【解決手段】制御プログラムP20は、ユーザ企業識別子 (#0等)ごとの計算機割当て制御表 (T22、T23)を作成し、負荷分散装置d100に設定する。ユーザ要求パケット中のIPアドレス (A0等)からユーザ企業識別子を識別する手段 (T19)を設ける。負荷分散装置d100は、ユーザ要求パケットからユーザ企業を識別し、しかるべき計算機群にそれを転送する。さらにP20はユーザごとのサービスレベル契約 (T20)と、計算機の稼動状況モニタ結果を比較 (T21中CPU使用率履歴)、契約条件未達時は割当て計算機を増やす。具体的にはT22、T23を変更し、負荷分散装置d100に再設定する。さらにP20は、それに基づく課金情報 (T21中サーバ割当て履歴、課金)を変更する。

図4

記憶ユニット

| | | | | | |
|------|-----|-----|--|--|--|
| addr | UID | T19 | | | |
| A0 | #0 | | | | |
| A1 | #0 | | | | |
| b0 | #1 | | | | |
| b1 | #1 | | | | |

| | | | | | |
|-----|----|----------|-----------|------|------|
| UID | Gr | CPU使用率履歴 | OUTスルーパート | 下限台数 | 上限台数 |
| #0 | AP | 50%未満 | — | 2台 | 8台 |
| #1 | DB | 50%未満 | — | 2台 | 8台 |
| web | — | 10件/秒以上 | — | 2台 | 6台 |
| #1 | : | : | : | : | : |

| | | | | | |
|--------|--------|----------|---------|----------|------|
| UID | サーバ | CPU使用率履歴 | ... | サービス割当履歴 | 課金 |
| #0 | web-Gr | 0.5時間 | 108.5時間 | 2台時間 | 80時間 |
| #1 | AP-Gr | 0時間 | 100時間 | 10時間 | 56時間 |
| #1 | DB-Gr | 0.5時間 | 108.5時間 | 20時間 | 60時間 |
| web-Gr | — | — | — | 10時間 | — |
| #1 | : | : | : | ... | 1.5M |

| | | | | |
|--------|----------|-------|----------|------|
| UID | CPU使用率履歴 | ... | サービス割当履歴 | 課金 |
| #0 | web-Gr | 0.5時間 | 108.5時間 | 2台時間 |
| #1 | AP-Gr | 0時間 | 100時間 | 10時間 |
| #1 | DB-Gr | 0.5時間 | 108.5時間 | 20時間 |
| web-Gr | — | — | — | 10時間 |
| #1 | : | : | : | ... |

| | | | | |
|------|---------|------|----------|-------|
| 制当ID | 制当IDサーバ | UID | 制当addr | 制addr |
| #0 | web-Gr | #100 | #10.a.11 | |
| #1 | AP-Gr | #200 | #20.a.21 | |
| #1 | DB-Gr | #300 | #30.a.31 | |
| #1 | web-Gr | #100 | #10.b.11 | |
| #1 | AP-Gr | #200 | #20.b.21 | |
| #1 | DB-Gr | #300 | #30.b.31 | |

制御プログラム P20

制御ユニット

1

【特許請求の範囲】

【請求項1】相互にネットワークで結合される複数の計算機を有し、複数ユーザの要求を処理し、要求ごとにそれがどのユーザに属するものかを識別するユーザ識別手段を含む計算機システムにおいて、各ユーザごとに異なる計算機群を割り当てる計算機資源分割方法であって、各要求からユーザ識別番号を検出するため必要な条件を入力する手順と、

該条件からユーザ識別表を作成する手順と、

該ユーザ識別表を、上記ユーザ識別手段に設定する手順と、

ユーザごとのサービスレベル条件を入力する手順と、該サービスレベル条件に基づきユーザ識別番号ごとに処理を実行すべき1つ以上の計算機の割当てを行なう手順と、

該割当てを定義する計算機割り当て定義表を生成する手順と含む計算機資源分割方法。

【請求項2】上記計算機システムは、上記複数の計算機に対して負荷を分散する手段をさらにふくみ、上記計算機割り当て定義表は該負荷分散手段に設定されることを特徴とする請求項1の計算資源分割方法。

【請求項3】相互にネットワークで結合される複数の計算機を有し、ユーザ識別子と処理依頼に際して使用することが指定された依頼先計算機識別子とがそれぞれ付加された複数ユーザの要求を処理する計算機システムに適用され、各計算機に処理を割り当てる負荷分散装置において、

ユーザごとにそのユーザからの上記依頼先識別子と依頼処理を実行すべき1つ以上の計算機を識別する識別子との対応を示す表を保持する手段、及び上記ユーザ要求を受け、該ユーザ要求から抽出したユーザ識別子と上記表とを参照して、該ユーザ要求に付加された依頼先計算機識別子をこれに対応する処理を実行すべき計算機の識別子のいずれかに変更し、該ユーザ要求を処理すべき計算機を特定して上記ユーザ要求を転送する手段を有する負荷分散装置。

【請求項4】上記ユーザ識別表は、ユーザ要求パケットの送り元IPアドレスにもとづきユーザを識別する表であることを特徴とする請求項1の計算資源分割方法。

【請求項5】上記ユーザ識別表は、ユーザ要求パケットの送り先IPアドレスにもとづきユーザを識別する表であることを特徴とする請求項1の計算資源分割方法。

【請求項6】相互に内部ネットワークで結合される複数の計算機を有し、外部ネットワークを介して接続された複数のユーザのシステムから入力する処理要求にそれぞれ応動して処理を行う計算機システムにおける計算機資源分割方法であり、

各ユーザの利用契約の締結の際に、該ユーザからの処理要求のアクセス先アドレスとして用いる仮想IPアドレスを該ユーザのユーザシステムに設定すること、該ユー

2

ザからの処理要求であることの識別に前記処理要求のアクセス元IPアドレスとアクセス先IPアドレスのいずれを用いるかを選択すること、及びその用いるIPアドレスを入力することを前記外部ネットワークを介して該ユーザに促し、

該ユーザからの処理要求について処理を実行すべき計算機の割り当て数の上限値、下限値を少なくとも含むサービスレベル条件の入力を前記外部ネットワークを介して前記ユーザに促し、入力されたサービスレベル条件に基づき前記ユーザの処理要求に対して処理を実行すべき計算機の割当てを行い、割り当てた計算機数の履歴を記録することを特徴とする計算機資源分割方法。

【請求項7】相互に内部ネットワークで結合される複数の計算機を有し、外部ネットワークを介して接続された複数のユーザのシステムから入力する処理要求にそれぞれ応動して処理を行う計算機システムにおける計算機資源分割方法であり、

各ユーザの利用契約の締結の際に、該ユーザからの処理要求のアクセス先アドレスとして用いる仮想IPアドレスを該ユーザのユーザシステムに設定すること、該ユーザからの処理要求であることの識別に前記処理要求のアクセス元IPアドレスとアクセス先IPアドレスのいずれを用いるかを選択すること、及びその用いるIPアドレスを入力することを前記外部ネットワークを介して該ユーザに促し、

該ユーザからの処理要求について処理を実行すべき計算機使用率を少なくとも含むサービスレベル条件の入力を前記外部ネットワークを介して前記ユーザに促し、入力されたサービスレベル条件に基づき前記ユーザの処理要求に対して処理を実行すべき計算機の割当てを行い、割り当てた計算機使用率の履歴を記録することを特徴とする計算機資源分割方法。

【請求項8】複数ユーザの要求を処理する計算機システムであって、

処理が割り当てられる相互にネットワークで結合された複数の計算機と、

複数のユーザの各自について、要求からユーザ識別番号を抽出する条件と、サービスレベル条件を受け、該条件からユーザ識別表を生成し、該サービスレベル条件に基づきユーザごとの計算機の割り当てを行って割り当て定義表を生成する管理手段と、

上記ユーザ識別表に従い、入力するユーザ要求からユーザ識別番号を抽出し、上記計算機割り当て表から要求を処理すべき計算機を特定して上記ユーザ要求を転送する付加分散手段とを有することを特徴とする計算機システム。

【請求項9】相互にネットワークで結合される1つ以上の計算機を有し、複数ユーザの要求を処理し、要求ごとにそれがどのユーザに属するものかを識別するユーザ識別手段を含む計算システムであって、各計算機では複数

50

3
 の異なるオペレーティングシステムが時分割で動作する
 よう構成され、各オペレーティングシステムは夫々専用
 に資源を使用し、また上記時分割動作の実行比率を定義
 できるよう構成されている計算機システムにおける各ユ
 ーザへの計算資源分割方法であって、
 各要求からユーザ識別番号を検出するためには必要な条件
 を入力する手順と、
 該条件からユーザ識別表を作成する手順と、
 該ユーザ識別表を、上記ユーザ識別手段に設定する手順
 と、
 ユーザごとのサービスレベル条件を入力する手順と、
 該サービスレベル条件に基づきユーザ識別番号ごとに処
 理を実行すべき計算機群の時分割実行比率表を生成する
 手順と、
 該実行比率を上記計算機に設定する手順とからなること
 を特徴とする計算資源分割方法。
 【請求項10】相互にネットワークで結合される1つ以
 上の計算機および計算機資源分割制御手段で構成され、
 複数ユーザの要求を処理する計算システムであって、
 上記計算機は、複数の異なるオペレーティングシステム
 が時分割で動作するよう構成され、各オペレーティング
 システムは夫々専用の資源を有するように構成され、また時
 分割動作する比率を定義できるよう構成されており、
 計算機資源分割制御手段は上記請求項7の計算資源
 分割方法に従って動作する手段であることを特徴とする
 計算システム上記計算機分割制御手段は、複数のユーザ
 の各々について要求からユーザ識別子を抽出する条件
 と、サービスレベル条件を受け付け、該条件からユーザ
 識別表を生成し、該サービスレベル条件に基づきユーザ
 ごとに処理を実行すべき計算機の時分割実行比率表を生
 成し、上記ユーザ識別表と時分割比率設定表を上記計算
 機に設定することを特徴とする計算機システム。
 【請求項11】相互にネットワークで結合される複数の
 計算機を有し、複数ユーザの要求を処理する計算システム
 において、各ユーザごとに計算機の割当てを自動的に
 変更する計算資源分割方法であって、
 上記計算機資源の稼動状況をモニタする手順と、
 上記稼動状況とユーザごとのサービスレベルを比較する
 手順と、
 上記比較に基づきユーザごとの計算機割当てを変更すべ
 きか判断する手順と、
 ユーザごとの計算機割当て表を変更する手順と、
 計算機割当て変更に基づき課金情報を変更する手順を有
 することを特徴とする計算資源分割方法。
 【請求項12】相互にネットワークで結合される複数の
 計算機で構成され、複数ユーザの要求を処理する計算機
 システムにおいて、各ユーザごとに計算機の割当てを自
 動的に変更する計算資源分割方法であって、
 上記計算機資源の稼動状況を受け取る手順と、
 上記稼動状況とユーザごとのサービスレベルを比較する

- 10 手順と、
 上記比較に基づきユーザごとの計算機割当てを変更すべ
 きか判断する手順と、
 変更が必要と判断されたときユーザごとの計算機割当て
 表を変更する手順を有することを特徴とする計算資源分
 割方法。
 【請求項13】上記計算機システムは複数の負荷分散手
 段をさらに含み、上記計算資源分割方法は、上記変更し
 たユーザごとの計算機割当て表を上記負荷分散手段に設
 定する手順と、上記複数の負荷分散手段の全てにおいて、
 設定が完了するまで待つ手順を更に含むことを特徴とす
 る請求項12の計算資源分割方法。
 【請求項14】上記複数の計算機は機能の異なる複数の
 計算機グループから構成され、上記計算機割当ては上記
 各計算機グループごとに同一グループに属する計算機の
 割当てであり、あるグループの計算機資源を増やす時は
 同一のグループに属する計算機から選択することを特徴
 とする請求項12の計算資源分割方法。
 【請求項15】相互にネットワークで結合される複数の
 計算機を有し、各計算機には標準でアクセスするルート
 ファイルが設定されており、複数ユーザの要求を処理す
 る計算システムにおいて、各ユーザごとに計算機の割當
 てを自動的に変更する計算資源分割方法であって、
 上記計算機資源の稼動状況を受け取る手順と、
 上記稼動状況とユーザごとのサービスレベルを比較す
 る手順と、
 上記比較に基づきユーザごとの計算機割当てを変更すべ
 きか判断する手順と、
 ユーザごとの計算機割当て表を変更する手順と、
 30 計算機ごとのルートファイル名を変更する指示を出す手
 順を有することを特徴とする計算資源分割方法。
 【請求項16】相互にネットワークで結合される複数の
 計算機および計算資源分割手段を有し、複数ユーザの要
 求を処理する計算機システムにおいて、上記計算資源分
 割手段は、上記計算機資源の稼動状況を受け取る手段
 と、上記稼動状況とユーザごとのサービスレベルを比較
 してユーザごとの計算機割当てを変更すべきか判断す
 る手段と、変更すべき場合にユーザごとの計算機割當
 て表を変更する手段とを有することを特徴とする計算機シ
 ステム。
 【請求項17】上記計算機資源分割手段は、上記計算機
 資源の稼動状況をモニタする手段と、計算機割当て変更
 に基づき課金情報を変更する手段をさらに有することを
 特徴とする請求項16記載の計算機システム。
 【請求項18】相互にネットワークで結合される1つない
 しは複数の計算機で構成され、複数ユーザの要求を処理す
 る計算システムであって、各計算機は複数の異なる
 オペレーティングシステムが時分割で動作するよう構成
 され、各オペレーティングシステムは夫々専用の資源を
 有するように構成され、また上記時分割動作する比率を
 50 有するように構成され、また上記時分割動作する比率を

定義できるよう構成されている計算機システムにおいて、各ユーザごとに計算機資源群割当てを自動的に変更する計算資源分割方法であって、
上記計算機資源群の稼動状況をモニタする手順と、
上記稼動状況とユーザごとのサービスレベルを比較する手順と、
上記比較に基づきユーザごとの上記時分割動作する比率を変更すべきか判断する手順と、
ユーザごとの時分割動作する比率表を変更する手順と、
時分割動作比率変更に基づき課金情報を変更する手順を有することを特徴とする計算資源分割方法
【請求項19】相互にネットワークで結合される1つ以上の計算機で構成され、複数ユーザの要求を処理する計算システムであって、該計算機は、複数の異なるオペレーティングシステムが時分割で動作するよう構成され、各オペレーティングシステムは夫々専用の資源を有するように構成され、また上記時分割動作する比率を定義できるよう構成されている計算機システムにおいて、各ユーザごとに上記計算機資源群割当てを自動的に変更する計算資源分割方法であって、
上記計算機資源群の稼動状況を受け取る手順と、
上記稼動状況とユーザごとのサービスレベルを比較する手順と、
上記比較に基づきユーザごとの上記時分割動作する比率を変更すべきか判断する手順と、
ユーザごとの時分割動作する比率表を変更する手順を有することを特徴とする計算資源分割方法。
【請求項20】相互にネットワークで結合される1つ以上の計算機および計算機資源分割制御手段で構成され、複数ユーザの要求を処理する計算システムにおいて、
上記計算機は、複数の異なるオペレーティングシステムが時分割で動作するよう構成され、各オペレーティングシステムは夫々専用の資源を有するように構成され、さらに上記時分割動作する比率を定義できるよう構成されており、
上記計算機資源分割制御手段は上記計算機群の稼動状況を受け取る手段と、上記稼動状況とサービスレベルを比較して上記時分割動作する比率を変更すべきか判断する手段と、変更すべき場合にユーザごとの上記時分割動作する比率を変更する手段とを有することを特徴とする計算機システム。

【請求項21】上記計算機資源分津制御手段は、上記計算機群の稼動状況をモニタする手段と、上記時分割動作する比率の変更にしたがって課金情報を変更する手段をさらに有することを特徴とする請求項20記載の計算機システム。

【請求項22】相互にネットワークで結合される複数の計算機および計算機資源分割制御手段を有し、複数ユーザの要求を処理する計算システムにおける各ユーザへの課金方法であって、

- 予め設定されたユーザごとのサービスレベルと計算機資源の稼動状況を比較して各ユーザの計算機割り当てを随時変更し、
ユーザ識別子ごとに割り当てた計算機数および時間を記録し、
上記計算機数と時間の積に基づき課金金額を算出することを特徴とする課金方法。
【請求項23】機能の異なる計算機グループにグループ分けされ、相互にネットワークで結合される複数の計算機、および計算機資源分割制御手段を有し、複数ユーザの要求を処理する計算システムにおける各ユーザへの課金方法であって、
予めユーザごとに設定されたサービスレベルと計算機資源の稼動状況との比較に応じて各ユーザの計算機割り当てを随時変更し、
ユーザ識別子ごとに上記グループごとの割り当て計算機数および時間を記録するし、
上記グループごとの計算機数と時間の積に基づき課金金額を算出することを特徴とする課金方法。
【請求項24】性能の異なる計算機グループにグループ分けされ、相互にネットワークで結合される複数の計算機、および計算機資源分割制御手段を有しで構成され、複数ユーザの要求を処理する計算システムにおける各ユーザへの課金方法であって、
予めユーザごとに設定されたサービスレベルと計算機資源の稼動状況との比較に応じて各ユーザの計算機割り当てを随時変更し、
ユーザ識別子ごとに上記グループごとの割り当て計算機数および時間を記録し、上記グループごとの計算機数と時間の積に基づき課金金額を算出することを特徴とする課金方法。
【請求項25】相互にネットワークで結合される複数の計算機および計算機資源分割制御手段を有し、複数ユーザの要求を処理する計算システムにおける各ユーザへの課金方法であって、
予めユーザごとに設定されたサービスレベルと計算機資源の稼動状況との比較に応じて各ユーザの計算機割り当てを随時変更し、
各ユーザから計算機システムに入力される要求数と計算機システムから各ユーザに返答する返答数を時間ごとに計測し、
計測結果に基づき課金金額を算出することを特徴とする課金方法。
【請求項26】相互にネットワークで結合される1つ以上の計算機および計算機資源分割制御手段を有し、複数ユーザの要求を処理する計算システムであって、上記計算機は、複数の異なるオペレーティングシステムが時分割で動作するよう構成され、各オペレーティングシステムは夫々専用の資源を有するように構成され、また上記時分割動作する比率を定義できるよう構成されている計

算システムにおける各ユーザへの課金方法であって、予めユーザごとに設定されたサービスレベルと計算機資源群の稼働状況との比較に応じて各ユーザへの計算機の時分割動作の時分割割り当て比率を随時変更し、ユーザ識別子ごとに割り当てた時分割割り当て比率、割り当て時間を記録し、上記割り当て比率と割り当て時間の積に基づき課金金額を算出することを特徴とする課金方法。

【請求項27】機能の異なる計算機グループにグループ分けされ、相互にネットワークで結合される複数の計算機、および計算機資源分割制御手段を有し、複数ユーザの要求を処理する計算システムであって、上記各計算機は、複数の異なるオペレーティングシステムが時分割で動作するよう構成され、各オペレーティングシステムは夫々専用の資源を有するように構成され、また上記時分割動作する比率を定義できるよう構成されている計算システムにおける各ユーザへの課金方法であって、

予めユーザごとに設定されたサービスレベルと計算機資源群の稼働状況との比較に応じて各ユーザへの計算機の割り当てと、時分割動作の時分割割り当て比率を随時変更し、

ユーザ識別子ごとに上記グループごとの割り当て計算機数、時分割割り当て比率、割り当て時間を記録し、上記グループごとの計算機数と割り当て比率と割り当て時間の積に基づき課金金額を算出することを特徴とする課金方法。

【請求項28】性能の異なる計算機グループにグループ分けされ、相互にネットワークで結合される複数の計算機、および計算機資源分割制御手段を有し、複数ユーザの要求を処理する計算システムであって、上記各計算機は、複数の異なるオペレーティングシステムが時分割で動作するよう構成され、各オペレーティングシステムは夫々専用の資源を有するように構成され、また上記時分割動作する比率を定義できるよう構成されている計算システムにおける各ユーザへの課金方法であって、予めユーザごとに設定されたサービスレベルと計算機資源群の稼働状況との比較に応じて各ユーザへの計算機の割り当てと、時分割動作の時分割割り当て比率を随時変更し、

ユーザ識別子ごとに上記グループごとの割り当て計算機数および時分割割り当て比率と割り当て時間を記録し、上記グループごとの計算機数と割り当て比率と割り当て時間の積に基づき課金金額を算出することを特徴とする課金方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、計算機システム内の資源をユーザごとに分割する機構および方法に係わり、特に相互にネットワークで結合される複数の計算機からなるシステムにおいて、複数ユーザの要求を処理す

る際に、各ユーザとの間であらかじめ取り決めたサービスレベルを維持するために必要な計算資源をリアルタイムに提供し、かつユーザ間のセキュリティを保持するための資源分割機構および方法に関する。

【0002】

【従来の技術】情報部門のコスト削減のために、企業内情報システム運用や企業のホームページ管理をASP(アプリケーション・サービス・プロバイダ)業者にアウトソースする事業形態が増加している。ASP業者はさらにデータセンタ業者に計算機資源供の供給、運用管理をアウトソースする形態が多い。

【0003】データセンタ業者は、多数の計算機資源を用意して複数のユーザ企業に分割使用させることにより、自身の運用コストを削減し、低価格のサービスをユーザ企業に提供する。ユーザ企業間のセキュリティを保持するため、一般には各ユーザ企業ごとに異なる計算機資源およびストレージ資源を割り当てることが多い。

【0004】またユーザ企業の負荷は時間帯、季節等により変動するため、負荷に応じて割り当て資源を増減させる

20 契約形態も多い。特に企業のホームページ管理を請け負っている場合などは、不特定多数の消費者がインターネット経由でアクセスするため負荷の予測がつきにくい。これに対してはユーザ企業側でたとえば新製品発表による負荷増を予測してあらかじめ決めた台数の計算機資源をある一定期間増強するようデータセンタ業者と契約する。データセンタ業者は、増強した計算機資源を、別の期間は別のユーザ企業に割り当てるにより資源の有効活用を図る。このような構成変更を容易にするために、負荷分散装置を複数の計算機資源の前段に配置

30 し、ある期間はユーザ企業Aに、別の期間はユーザ企業Bに計算機資源を使用するように、データセンタを構成する。負荷分散装置の例としては、アルテオン社のACE directorなどが挙げられる(日経オーブンシステムズ1999.12 no.81 pp.128-131)。負荷分散装置の設定は上記の契約に基づきあらかじめ人手で設定する。またストレージ資源を増強する必要がある場合は、ストレージの内容をミラーリングする必要がある。

【0005】さらに、データセンタでは多数のユーザ企業に異なる計算機資源を提供するため、多数の計算機資源を管理する必要が生じ管理コストが増大する。そこで1台あたりの性能が高い計算機資源、たとえば高多重SMP計算機を数少なく導入し、それを複数ユーザ企業が共有するように制御する方法が考えられる。ユーザ企業間のセキュリティを保持するため仮想計算機の機能を利用する。仮想計算機の例としては、日立のプロセッサ資源分割管理機構PRMFがあげられる(HITACマニュアル8080-2-148-60)。PRMFでは、1つの計算機上で複数のOS(オペレーティング・システム)が動作し、OSごとに独立の資源、たとえば

50

主記憶、ネットワークアダプタなどが割り当てられる。OS間で資源を共有しないため、異なるOS上で実行される異なるユーザ企業のプログラム間のセキュリティは保持される。またPRMFでは、OSごとに割り当てるCPU資源の比率を制御できるように構成されているが、あらかじめ計画した比率変更のみ可能である。

【0006】ASP、ISP（インターネット・サービス・プロバイダ）業者とユーザの間ではサービスレベル契約を結ぶことが一般的になりつつある。接続性、可用性、レイテンシ性能などのサービスレベル保証を契約する。さらに保証レベル未達の場合の補償契約を結ぶ形態も多い。

【0007】

【発明が解決しようとする課題】上記の従来技術では、以下の問題点がある。

【0008】負荷分散装置の設定を契約に基づきあらかじめ人手で設定する方式では、ユーザ企業側で予測できない急速な負荷変動にリアルタイムに対応することは困難である。これはユーザ間で異なる計算機を割り当てる場合も仮想計算機を使用する場合も同様である。またストレージ資源を増強する場合、ミラーリングに伴うデータコピーのオーバヘッドにより迅速な増強が困難である。さらにデータセンタの場合処理内容が定型でなく、ひとつのユーザからの処理要求が複数の計算機にまたがって処理される場合など、レイテンシ性能などは定義、測定しにくい。

【0009】

【課題を解決するための手段】上記を解決するため、本発明では、データセンタの計算機資源およびストレージ資源をユーザ負荷変動に基づきリアルタイムにユーザ企業ごとに分割割当する資源分割手段および方法を提供する。

【0010】そのため本発明では、ユーザ識別表を用意し、データセンタへのユーザ要求パケットからユーザ企業を特定し、パケットにユーザ識別子を付加する。さらにユーザ識別子ごとに処理を実行すべき計算機の組を定義する表を管理サーバで生成し、本表を負荷分散装置に動的に設定する。負荷分散装置は設定された計算機の組からいざれかを選択してユーザ要求を実行させる。複数の負荷分散装置がある場合は、管理サーバは負荷分散装置間で本表が整合するよう制御する。さらに、管理サーバでは各計算機の稼動状況をモニタし、サービスレベル契約を満たしているか調べ、必要であれば計算機資源の削減・増強を行う。具体的には上記の計算機の組の定義表を変更して、負荷分散装置に再設定する。さらに管理サーバはユーザ識別子対応に割当てた計算機資源量やサービスレベル契約を守れたかの履歴を作成し、課金情報を作成する。またデータセンタ全体の処理スループットを計測するために、データセンタに入力されるユーザ要求数と返答数をユーザ識別子ごとに測定し集計する。

【0011】また本発明の別の実施形態では、仮想計算機構造を有する計算機でデータセンタを構成する。ユーザ企業ごとに1つのOSで制御される仮想計算機構造を与え、管理サーバは各計算機に各計算機のCPU時分割使用割当%を動的に設定する。さらに、管理サーバでは各計算機の稼動状況をモニタし、サービスレベル契約を満たしているか調べ、必要であればCPU時分割割当%の削減・増強を行う。

【0012】

10 10 【発明の実施の形態】以下、図面を参照しながら、本発明の実施の形態を説明する。

【0013】(1) 第1の実施形態

図1は、本発明で対象とするデータセンタが、インターネット（IIO）を介してユーザ企業A（AA0）、ユーザ企業B（BB0）、A社やB社のホームページにアクセスする一般消費者（c0、c1）と接続されている例を示す。ここでクライアントa0、a1、a2はA社システムのプライベート・ネットワーク・アドレスを有しており、ゲートウェイA0、A1を介してVPN（パーティクルプライベートネットワーク）経由でデータセンタ内のゲートウェイD0と接続する。なお、クライアントc0、c1からの要求に関しては第3の実施形態において述べる。

【0014】図2は、データセンタDD0の構成を示す。本例ではユーザ要求に対しwebブラウザインターフェースを提供するwebサーバ群、webサーバを経由して発生するアプリケーションプログラムを動作させるAPサーバ群、アプリケーションプログラムからデータベースアクセス要求が出たときにそれを処理するDBサーバ群の3層構成になっている場合を示す。

【0015】図22は、ユーザ企業Aが本データセンタと使用条件契約を結ぶときの入力画面の一例を示している。本例では、ゲートウェイD0に入力されるアクセス要求が、ユーザ企業Aに属するユーザからのアクセス要求であることを識別するために、要求パケットのアクセス元IPアドレスがA0またはA1であることを利用する契約を結んでいる。またユーザ企業Aはデータセンタ内のwebサーバ群、APサーバ群、DBサーバ群のすべてを使用し、ユーザ企業Aのユーザ要求により起動されるプログラムは、webサーバのIPアドレスとしてa100、APサーバのIPアドレスとしてa200、DBサーバのIPアドレスとしてa300を使用する契約を結んでいる。

【0016】図23は、ユーザ企業Aが本データセンタとサービスレベル契約を結ぶときの入力画面の一例を示している。本例では、ユーザ企業Aのために、webサーバ、APサーバ、DBサーバいずれも必ず最低2台を割当て、またそのすべてをCPU稼働率50%未満で稼動させる。稼働率が50%以上となる場合は各webサーバ、APサーバ、DBサーバの割当台数を最大8台ま

では増加する契約を結んでいる。本例では入力画面中にチェック印が入力されていないが、たとえばデータセンタの出口における出力トランザクションのスループット、出力トランザクションと入力トランザクションのスループット比、トランザクション処理のレイテンシをサービスレベル契約とともに可能である。

【0017】上記のような入力画面による契約に基づき、A社にはwebサーバa10, a11, APサーバa20, a21, DBサーバa30, 31、が与えられ、B社にはwebサーバb10, b11, APサーバb20, b21, DBサーバb30, b31が与えられるものとする。またストレージS1はボリューム単位にA社、B社に割り当てられる。ボリュームV0がA社用、V1がB社用である。ストレージS2, S2も同様に割り当てられているが記述を省略している。サーバ群y10～y31は、A社、B社の負荷が大きくなつたときに割り当てるための予備サーバ群である。

【0018】A社が使用するIPアドレスは、webサーバがa100、APサーバがa200、DBサーバがa300とする。またB社も同様の入力画面により、webサーバb100、APサーバb200、DBサーバb300というIPアドレスを使用するものとする。

【0019】以下図を参照しながら、サーバ群および負荷分散装置が、ユーザAの要求をサーバ群a0～a31を使用して如何に処理するかを示す。

【0020】図1においてクライアントa0がゲートウェイA0に送出した要求パケットの構成は、図7(1)1200に示すとおりである。パケットの先頭があて先サーバのアドレス、次項が送り元クライアントのアドレスである。ゲートウェイA0はパケットをインターネットI10に送出するときにVPNのカプセル化を行い、図7(2)のパケット1201を生成する。ゲートウェイD0ではこのパケットをアンカプセル化する。本技術は公知である。

【0021】図3はデータセンタの入口にあるゲートウェイD0の構成図である。ゲートウェイD0は信号線I0から入力された図7(2)のパケットに対しアンカプセル化を行うとともに、ユーザ識別表T10を参照してユーザ識別子#0を得、パケットに#0を付加し、図7(3)のパケット1202を生成して信号線L10に送出する。ユーザ識別表T10は、あらかじめ図22のユーザ条件入力画面に基づき管理サーバC0に生成され、信号線L0を介して設定されている。すなわち、本データセンタに送り元アドレスA0またはA1でアクセスした要求はユーザ識別子#0、すなわちA社ユーザとみなされる。

【0022】ゲートウェイD0はパケット生成と同時に、入出力結果格納表T11に、ユーザ識別子#0の入力要求が通過したことを計数手段1003を用いて積算して記録する。

【0023】パケット1203を信号線L10を介して受け取った負荷分散装置d100は図5(1)に示すサーバアドレス対応表T30が保持されている。本表は、ユーザ識別子ごとにユーザアプリケーションで認識しているサーバへの要求を、どの実サーバに送出すべきかを記録しているものである。パケット1202のユーザ識別子は#0、あて先a100なので、負荷分散装置d100は表T30を参照してあて先サーバをa10かa11のいずれかに変更し、図7(4)のパケット1203を生成する。あて先の選択および変更技術自身は公知である。

【0024】webサーバa10はパケット1203を受け取り、a200へのアクセス要求であるパケット1204(図7(5))を生成する。本パケットはバスL110を介して負荷分散装置d200に受け取られる。負荷分散装置d200は図5(2)のサーバアドレス対応表T31を保持しており、それに従い、あて先サーバa200をa20に変更したパケット1205(図7(6))を生成する。

【0025】以下同様にAPサーバa20がパケット1206を生成し、サーバアドレス対応表T32(図5(3))を保持する負荷分散装置d300がパケット1206をパケット1207に変更し、DBサーバa30で処理が行われる。DBサーバa30からAPサーバa20、webサーバa10、クライアントa0に返答が返るのも同様であり、順次パケット1208～1214が生成される。ゲートウェイD0は返答パケットをゲートウェイA0に対して送出するときに、入出力結果格納表T11に、ユーザ識別子#0の出力要求が通過したことを見計数手段1003を用いて積算して記録する。

【0026】図示していないが、ユーザ企業Bからの要求があった場合は、上記と同じ手順でゲートウェイD0でユーザ識別子#1が付加され、同様にサーバ群b10～b31で処理されることとなる。

【0027】以上により、ユーザAとBの処理を実行するサーバ群はサーバa10～a31、b10～b31に夫々分割される。

【0028】ストレージへのアクセスについては図2のストレージS0を例に説明する。ストレージS0は信号線L120によりすべてのwebサーバにより共有されている。ただし、各サーバはストレージアクセス時にユーザ識別子をアクセス要求に付加する。一方ストレージS0は図6のボリュームアクセス権限表T33を保持している。本表はユーザ識別子ごとにどのボリュームのアクセスが許可されているかを示したものである。もしユーザ識別子#1のアクセス要求がボリュームV0をアクセスしようとした場合、ストレージS0は本表を参照し、アクセスを拒否する。これにより、ストレージS1がすべてのwebサーバに共有されていても、ユーザAとBの間のセキュリティが保たれる。

【0029】図2において、管理サーバC0は信号線L100, L200, L300を介してサーバ群および負荷分散装置の動作状況をモニタリングしている。モニタリングの内容はユーザとのサービスレベル契約の内容や、モニタリングプログラムの機能による。たとえばCPU稼働率や負荷分散先履歴などである。またモニタリングプログラムは管理サーバC0で動作してもよいし、各サーバ群、負荷分散装置上で動作してもよい。また管理サーバC0は信号線L0を介してゲートウェイD0からユーザごとの入出力結果表T11の内容を入手する。

【0030】図4は管理サーバC0の構成図である。T19はユーザ識別表であり、制御プログラムP20が図22のユーザ条件入力画面に基づき設定する。T20はユーザごとのサービスレベル契約内容表であり、制御プログラムP20が図23のサービスレベル条件入力画面に基づき設定する。この場合、ユーザ識別子#0のユーザに対してはwebサーバ、APサーバ、DBサーバいずれも最低2台を与え、与えたすべてのサーバでCPU稼働率50%未満でプログラムを動作させ、それに違反しそうな場合は8台まではサーバ数を増加する契約となっている。またユーザ識別子#1のユーザに対しては、webサーバ、APサーバ、DBサーバいずれも最低2台を与え、データセンタからのアクセス返答スループットは秒あたり30件以上を維持し、それに違反しそうな場合は6台まではサーバ数を増加する契約となっている。制御プログラムP20はモニタリング結果とサービスレベル契約内容表T20を照合し、現在の資源割当がサービスレベル契約を満たしているかを調べ、その結果をサービス履歴格納表T21に格納する。サービス履歴格納表T21には、たとえばユーザ識別子#0に与えたすべてのサーバのCPU稼働率履歴を記録する。制御プログラムP20はモニタリング結果がサービスレベル契約を満たしていない場合は、割当てサーバを増やす。そのためどのユーザにどのサーバを与えたかを示すサーバ割当て管理表T22や、ユーザアブリケーションで認識しているサーバ名と与えた実サーバの対応表であるサーバアドレス対応表T23を保持している。T23は各負荷分散装置d100～d300が保持するサーバアドレス対応表T30～T32のマスタである。またサーバ履歴格納表には課金情報もあわせて記録している。図示していないが、もしユーザとの契約が割当てサーバ数に応じて課金を増やすものであれば、課金計算式が変わるのでそれを反映する。またもしユーザとの契約が契約したサービスレベルを維持できない割合に応じて課金が変動するものであればやはりそれを反映する。

【0031】上記制御を行うために、制御プログラムP20が最初に資源を分割する手順を図20を用いて説明する。

【0032】最初に図22のユーザ条件入力画面に示される情報を入力し、ユーザ識別表T19を作成する(1

901)。つづいてその情報を信号線L0を介してゲートウェイD0にも設定する(1902)。

【0033】さらに図23のサービスレベル条件入力画面に示される情報を入力し、サービスレベル契約内容表T20、およびサーバアドレス対応表T23中の仮想add欄を作成する(1903)。続いてサーバの割当てを、webサーバ群、APサーバ群、DBサーバ群ごとに行う。具体的には、サービスレベル契約内容表T20を参照して各ユーザに最低2台ずつサーバを与えるべきことを検出すると、サーバ割当て管理表T22およびサーバアドレス対応表T23の実add欄を作成する(1904)。続いて作成したサーバアドレス対応表T23の必要な部分のコピーを信号線L100, 200, 300を介して負荷分散装置d100, 200, 300に設定する(1905)。

【0034】さらに、サービスレベル契約内容表T23に基づき、サービス履歴格納表T21を作成する。具体的には、ユーザ#0にはCPU稼働率履歴を記録する欄を、ユーザ#1にはトランザクション出力スループット履歴(図示せず)を記録する欄を作成する。

【0035】以上により資源分割制御に必要な情報が生成され、ゲートウェイD0、負荷分散装置d100, 200, 300に設定され、正しく資源分割された状態でシステムが動作開始できる。

【0036】続いて以下に、制御プログラムP20が負荷増大時に資源割当てを変更する手順を図8を用いて説明する。

【0037】前述したように信号線L100, 200, 300, 0を介してシステムの稼動情報をモニタし(1301)、ユーザ識別子ごとに稼動情報を集計してサービス履歴格納表T21に格納し(1302)、サービスレベル契約内容表T20と比較した後(1303)、まずサービスレベル契約に照らしてサーバを削減できないか検討する(1304)。削減可能かどうかの判断方法としては、CPU稼働率とサーバ台数の積に対して比例計算を行う方法が挙げられる。たとえばユーザ#0のサービスレベル条件はCPU稼働率50%未満であるが、現在4台がwebサーバとして与えられており、いずれもCPU稼働率が25%未満であれば、単純な比例計算としてはwebサーバ数を2台まで削減してよいと判断できる。実際はこれに経験から与えられる種々の安全計数を掛けて判断する。削減可能であれば、削減対象であるサーバへの処理停止指示を信号線L100, 200, 300のいずれかを介して通知する。通知されたサーバはプログラムの処理を終了して、使用している資源を解放する。すなわちメモリアドレス変換テーブルやキャッシュ内容の無効化などを行う。解放が終了するとサーバは管理サーバにその旨通知するので、管理サーバはそれを待って、負荷分散装置d100～d300にサーバアドレス対応表T30～T32の変更を指示する。続いて

すべての負荷分散装置の内容が一致して変更されたことを確認する。さらに課金計算式を変更する。本例の場合、割当てサーバ数と割り当てた時間の履歴を記録している。課金計算では、単位時間あたりの1サーバ割当てあたりに単価を決めておきそれを課金する。すなわち、総割当てサーバ数と割当て時間履歴と単価を乗じて課金を計算する(1305)。なお、本例では、webサーバ群、APサーバ群、DBサーバ群の割当て履歴を区別して記録しているので、群ごとに単価を変えておき群ごとの割当て数と割当て時間履歴と各単価の積で課金を計算することも可能である。また、本例では図示していないが、サーバごとに実効性能が異なる場合は、サーバ数×実効性能×割当て時間履歴×単価で課金計算できることも明らかである。また、本例ではゲートウェイD0を要求パケットが通過した個数と返答パケットが通過した個数を記録しているが、要求パケットのゲートウェイ通過スループットが比較的安定している時は、返答パケットのゲートウェイ通過スループットがデータセンタ処理能力の目安となる。そこで信号線L0を介してゲートウェイから返答パケットの通過スループットを受け取り、予め契約で定めた基準スループットと比較して課金を計算することも可能である。例えば、基準スループットを満たした時間は規定課金を行ない、満たさなかった時間は違約金分を課金から減ずることができる。また基準スループットの単価を決めておき、(計測スループット/基準スループット×単価)で課金を計算することもできる。また要求パケットの入力スループットが大きく変動する場合は、(返答パケットスループット/要求スループット)に基づき課金を行なう方法もある。

【0038】図8の説明に戻ると、つづいてサーバ数を増強する必要があるかを検討する(1306)。何台増強すべきかの判断は、削減時と同様に比例計算で行えばよい。増強する必要があれば、webサーバ、APサーバ、DBサーバ群ごとに割り当てられる空きサーバがあるかをサーバ割当て管理表T22を参照して調査する(1307)。もし空きサーバがあれば運用管理者に通知する(1308)。空きサーバがあれば割当てサーバを選択し(1309)、負荷分散装置d100～d300にサーバアドレス対応表T30～T32の変更を指示する。すべての負荷分散装置の内容が一致して変更されたことを確認してから、課金計算式を変更する(1310)。

【0039】以上は管理サーバCO上の制御プログラムP20の手順の一例であるが、必ずしもこのすべてをP20が実施しなくともよいことは明白である。たとえば稼動情報の収集や集計は本プログラムが行わずに他のプログラムからの受け取ってもよい。また本プログラムがかならず実施すべき1305、1310の処理の内容についても、図9の(1)に示すように課金情報を変更しない1401、1402に夫々置き換えてよい。また

サーバ側に処理停止指示後は新しいユーザ要求を受け付けない機能がある場合などは処理停止完了を待たずにサーバアドレス対応表T30～T32を変更を指示する1403、1404に夫々置き換えてよい。

【0040】なお上記ではストレージ資源のボリュームアクセス権限表を変更していないが、各プログラムはユーザ識別子を付加してストレージにアクセスするので、サーバ割当てが変更されてもアクセス権限のないボリュームにアクセスすることは避けられる。

10 【0041】(2) 第2の実施形態

続いて仮想計算機機能PRMFを備えた高多重SMPサーバを用いてデータセンタを構成する場合の実施形態を説明する。

【0042】データセンタとユーザの接続図は図1に同じである。

【0043】図10は、webサーバ、APサーバ、DBサーバとして仮想計算機機能PRMFを備えたサーバをおののおの1台ずつ接続した場合を示している。APサーバ1501、DBサーバ1502の内部構成はwebサーバ1500と同様であり記述省略する。

【0044】ユーザ条件入力画面は図22と同じである。すなわち、ユーザ要求パケットの送り元IPアドレスがA0、A1である要求のみがユーザ企業Aのパケットとみなされる契約である。またユーザ企業Aが使用するIPアドレスはwebサーバがa100、APサーバがa200、DBサーバがa300である。

【0045】図24は、サービスレベル契約条件入力画面の一例である。本例では、ユーザ企業Aのために、webサーバ、APサーバ、DBサーバいずれも必ずPR

30 MF機能によるCPU割当てが50%以上となるよう制御する契約である。

【0046】図10に戻ると、webサーバ1500は、制御ユニット1503、LPAR制御レジスタ1504、CPU群1505および1506、メモリ1507、ネットワークアダプタa100、b100、y100から構成される。LPARとはLogical Partition(論理資源分割)の略であり、LPAR制御レジスタ群1504は、各OSに与える資源の分割方法を保持している。

40 【0047】図11はLPAR制御レジスタ群1504が保持する情報の例である。従来技術であるPRMFでは、本情報のうちUIDの欄以外の情報を有している。

LPAR#とは、各OSごとに与える資源群に統一的に与える識別子である。ネットワークアダプタはLPARごとにあたえる。ネットワークアダプタアドレスは、後述するが制御プログラムP20により、ユーザ条件入力画面で契約した各ユーザに与えられるIPアドレスと一致するよう設定される。これによりあるネットワークアダプタに入ってきたユーザ要求のパケットは、対応する

50 LPARのOS上のプログラムに引き渡される。メモリ

割当て欄は、各 LPAR がメモリ 1507 のどの部分を使用するかを示す情報である。CPU 割当て % 欄は各 LPAR に属する OS およびその上のプログラムをどのような割合で CPU 上で動作させるかを示す。制御ユニット 1503 が本情報を参照して、LPAR の動作割合を制御する。

【0048】本実施例では UID 欄を追加し、LPAR とユーザ識別子を一意に対応させる。PRMF の制御により、異なる LPAR 間では資源が共有されないので、ユーザ間のセキュリティを保つことができる。

【0049】第 1 の実施形態と同様、ユーザ要求がクライアント a0 → web サーバ a100 → AP サーバ a200 → DB サーバ a300 → AP サーバ a200 → web サーバ a100 → クライアント a0 と伝わる場合を考える。クライアント a0 は図 12 (1) のパケット 1200 を生成する。ゲートウェイ A0 によりパケット 1201 が生成され、ゲートウェイ D0 によりパケット 1202 が生成されるのは第 1 の実施形態と同じである。

【0050】パケット 1202 は信号線 L0 を介してアドレス a100 を持つネットワークアダプタ a100 に渡され、LPAR #0 上のアプリケーションプログラムに渡される。すなわちユーザ A のアプリケーションプログラムである。本プログラムは先に a200 を持つパケット 1204 を生成し、以下同様に AP サーバ 1501 上の A 社のアプリケーションプログラム、DB サーバ 1502 上の A 社のアプリケーションプログラムに渡される。(図示していないが、AP サーバ 1501 内にはネットワークアダプタ a200、b200、y200 があり、夫々 LPAR #0, 1, 2 に対応している。また LPAR #0, 1 は夫々ユーザ識別子 #0, 1 に対応している。DB サーバ 1502 においても同様である。)

同様にして DB サーバ 1502 から AP サーバ 1501、web サーバ 1501、クライアント a0 への返答も正しく A 社に割り当てられた LPAR 上のアプリケーションプログラムにより実施される。詳しくは説明しないが、上記の動作により図 12 のパケット 1206 ～ 214 が順次生成される。

【0051】図 13 は管理サーバ C0 の構成図である。T40 は LPAR 割当て管理表、T19 はユーザ識別表である。T50 はユーザごとのサービスレベル契約内容表である。この場合、ユーザ識別子 #0 のユーザに対しては web サーバ、AP サーバ、DB サーバのいずれにおいてもユーザ識別子 #0 の LPAR を CPU 割当て率を 50% 以上割り当てる契約となっている。またユーザ識別子 #1 のユーザに対しては、CPU 割当て % を最低 20% とし、またデータセンタからのアクセス返答スループットは秒あたり 30 件以上を維持し、それに違反しそうな場合は CPU 割当て % を増加する契約となっている。制御プログラム P20 は信号線 L100, 200, 300, 0 から得たモニタリング結果とサービスレ

ベル契約内容表 T50 を照合し、現在の資源割当てがサービスレベル契約を満たしているかを調べ、その結果をサービス履歴格納表 T51 に格納する。たとえばユーザ識別子 #0 に対応する LPAR の実際の CPU 使用率履歴を記録する。またユーザ識別子 #1 でアクセス返答スループットが秒あたり 30 件未満の場合は、設定 CPU 割当て % を増やす。そのためどのユーザにどれだけ CPU 割当てを設定しているかを示す CPU 割当て管理表 T52 保持している。T52 は各 web サーバ、AP サーバ、DB サーバ内の LPAR 制御レジスタ群の CPU 割当て % 欄と同じの内容を保持している。サービス履歴格納表 T51 の課金情報欄の操作は第 1 の実施形態と同様である。

【0052】上記制御を行うために、制御プログラム P20 が最初に資源を分割する手順を図 21 を用いて説明する。

【0053】最初に図 22 のユーザ条件入力画面に示される情報を入力し、ユーザ識別表 T19 を作成する (2001)。つづいてその情報を信号線 L0 を介してゲートウェイ D0 にも設定する (2002)。

【0054】さらに図 24 のサービスレベル条件入力画面に示される情報を入力し、サービスレベル契約内容表 T50、および LPAR 割当て管理表 T40 中のネットワークアダプタ欄を作成する (2003)。

【0055】統いてサービスレベル契約内容表 T50 を参照してユーザ #0 に最低 50%、ユーザ #1 に最低 20% の CPU 割当てを行なべきことを検出すると、CPU 割当て管理表 T52 をと LPAR 割当て管理表 T40 中の CPU 割当て欄を作成する (2004)。さらに信号線 L100, 200, 300 を介してサーバ 1500, 1501, 1502 内の LPAR 制御レジスタ群に LPAR 割当て管理表 T40 の内容を設定する (2005)。さらに、サービスレベル契約内容表 T23 に基づき、サービス履歴格納表 T21 を作成する (2006)。

【0056】以上により資源分割制御に必要な情報が生成され、ゲートウェイ D0、サーバ 1500, 1501, 1502 に設定され、正しく資源分割された状態でシステムが動作開始できる。

【0057】統いて以下に、制御プログラム P20 が負荷増大時に資源割当てを変更する手順を図 14 を用いて説明する。

【0058】稼動情報収集 (1601)、稼動情報集計 (1602)、サービスレベル契約との比較 (1603) は第 1 の実施形態と同様である。その後 CPU 割当て率を削可能か検討し (1604)、削減可能であれば該当するサーバの LPAR 制御レジスタ群の内容を変更するよう指示する。削減可能化の判断方法は第 1 の実施形態と同様である。変更終了を待った後、課金計算式を変更する (1605)。本例の場合、割当て CPU 使用率

と割り当てた時間の履歴を記録している。課金計算では、webサーバ、APサーバ、DBサーバそれぞれについて単位時間あたりの使用単価を決めておき、単価×CPU使用率を合計して課金する。もちろん、webサーバ、APサーバ、DBサーバそれぞれの単価を異なるように設定してもよく、また単価をサーバの実効性能によって定めてよい。

【0059】つづいてCPU割当率を増強すべきか検討し(1606)、増強する必要があれば、該当するサーバに設定しているサーバ割当率の合計が100%を超えないか調べる(1607)。超えるならば運用管理者に通知する(1608)。超えなければ該当するサーバのLPAR制御レジスタ群の内容を変更するよう指示し、変更終了を待った後、課金情報を変更する(1609)。

【0060】(3) 第3の実施形態

最後に、不特定多数の一般消費者が企業Aや企業Bの提供するホームページにアクセスする場合の実施形態を説明する。

【0061】データセンタとユーザの接続図は図1に同じである。ユーザはクライアントc0、c1である。

【0062】図15はデータセンタの構成を示す。第1の実施形態と同じく、負荷分散装置d100により負荷を複数のサーバに分散できる形態とする。説明の簡単化のため、webサーバしかない構成を示す。すべてのwebサーバは信号線L120を介してストレージS4を共有している。S4には企業Aのホームページ情報を含むファイルF0と企業Bのホームページ情報を含むファイルF1が格納されている。ホームページ情報はトリー構成になっていて、ルートのページから順次たどれる構成とする。また企業Aの提供するホームページにアクセスするためのアドレスがa100、企業Bのためのアドレスがb100であるとする。

【0063】図25は企業Aのホームページにアクセスするユーザの条件を契約するための入力画面の一例である。本例では、ゲートウェイD0に入力されるアクセス要求が、企業Aのホームページにアクセスするユーザ群を識別するために、要求パケットのアクセス先IPアドレスがa100であることを利用する契約を結んでいる。また企業Aのホームページ生成のために使用されるIPアドレスはa100である契約を結んでいる。

【0064】クライアントc0は企業Aのホームページにアクセスするため図16(1)のパケット1700を生成する。ゲートウェイD0は図17に示すようにユーザ識別表T60を保持しており、パケット1700の先アドレスがa100であることから本パケットがユーザ識別子#0のホームページにアクセスしようとしていることを検知し、図16(3)のパケット1702を生成する。以下負荷分散装置d100により本アクセス要求はwebサーバa10またはa11のいずれかに送

られる。この場合はサーバa10が選択され、パケット1703(図16(4))が生成される。返答パケットも同様に、負荷分散装置d100によりパケット1712に変更され、さらにゲートウェイD0によるパケット1714に変形され、クライアントc0へと返る。

【0065】管理サーバC0内の構成は図18に示す。ルートファイル管理表T70が追加されている以外は図4と同じである。本表は、ユーザ識別子ごとのホームページのルートページのファイル名を保持している。

10 【0066】負荷が増大したときの制御プログラムP20の手順を図19に示す。本手順は図8とほぼ同じである。図8との相違である手順1800のみを説明する。手順1309において割当てるサーバを選択すると、選択したサーバに対して、ルートファイル管理表T70を参照して、これから割り当てるユーザ識別子に対応するルートファイル名を登録するよう指示する。その後は図4の手順1310と同じく、負荷分散装置d100にサーバアドレス対応表T30の変更を指示し、変更終了を待ってから、課金情報を変更する。ルートファイル名を変更することにより、新たに割当されたサーバは正しいホームページをアクセスすることができるようになる。

20 【0067】
【発明の効果】以上述べたように、本発明ではユーザ企業ごとにユーザ識別子を与え、それに基づき計算資源を与えるとともに、計算機の稼動状況のモニタ結果に基づき自動的にユーザ識別子ごとにサービスレベル契約と比較して計算資源の量を増減できる。これにより、ユーザ企業側で予測できない急速な負荷変動に対してもリアルタイムに計算資源の割当を変更可能である。

30 【0068】また計算機資源割当を変更しても、ストレージ資源を全計算機で共有し、かつユーザ識別子に基づきストレージ側でアクセス権限チェックを行っているため、ミラーリングのオーバヘッドなしにユーザ間セキュリティを維持することができる。

【0069】さらにデータセンタに入力される要求と返答の単位時間あたり通過数をユーザ識別子ごとに測定、集計するため、ユーザから見たデータセンタの性能を測定しやすい。

【図面の簡単な説明】
40 【図1】インターネットを介して接続されるデータセンタとユーザの例である。

【図2】データセンタの構成の例である。
【図3】図2のゲートウェイの構成図である。
【図4】図2の管理サーバの構成図である。
【図5】図2の負荷分散装置の保持するテーブル例である。

【図6】図2のストレージが保持するテーブル例である。

【図7】図2の信号線を通過するパケットの構成図である。

【図8】図4の制御プログラムの定常動作フロー例である。

【図9】図4の制御プログラムの定常動作フローの別の例である。

【図10】データセンタの構成の別の例である。

【図11】図9のL P A R制御レジスタ群の保持情報である。

【図12】図9の信号線を通過するパケットの構成図である。

【図13】図9の管理サーバの構成図である。

【図14】図13の制御プログラムの定常動作フロー例である。

【図15】データセンタの構成のさらに別の例である。

【図16】図15の信号線を通過するパケットの構成図である。

【図17】図15のゲートウェイの構成図である。

【図18】図15の管理サーバの構成図である。

【図19】図18の制御プログラムの動作フロー例である。

【図20】図4の制御プログラムの初期動作フロー例である。

【図21】図13の制御プログラムの初期動作フロー例である。

【図22】図2のデータセンタを使用するユーザ条件入力画面。

【図23】図2のデータセンタを使用するユーザのサービスレベル条件入力画面である。

【図24】図9のデータセンタを使用するユーザのサービスレベル条件入力画面である。

【図25】図15のデータセンタを使用するユーザ条件

*30

【符号の説明】

D D 0 : データセンタ、

D 0 : ゲートウェイ、

d 1 0 0 ~ d 3 0 0 : 負荷分散装置、

a 1 0 ~ a 3 1 、 b 1 0 ~ b 3 1 、 y 1 0 ~ y 3 1 : サーバ、

C 0 : 管理サーバ、

S 0 ~ S 4 : ストレージ、

10 T 1 0 、 T 1 9 : ユーザ識別表、

T 1 1 : 入出力結果格納表、

1 0 0 3 : 計数手段、

T 2 0 : サービスレベル契約内容表、

T 2 1 : サービス履歴格納表、

T 2 2 : サーバ割当て管理表、

T 2 3 、 T 3 0 ~ T 3 2 : サーバアドレス対応表、

T 3 3 : ボリュームアクセス権限表、

1 2 0 0 ~ 1 2 1 4 : パケット、

1 5 0 0 ~ 1 5 0 2 : サーバ、

1 5 0 4 : L P A R制御レジスタ、

T 4 0 : L P A R割当て管理表、

T 5 0 : サービスレベル契約内容表、

T 5 1 : サービス履歴格納表、

T 5 2 : C P U割当て管理表、

1 7 0 0 ~ 1 7 1 4 : パケット、

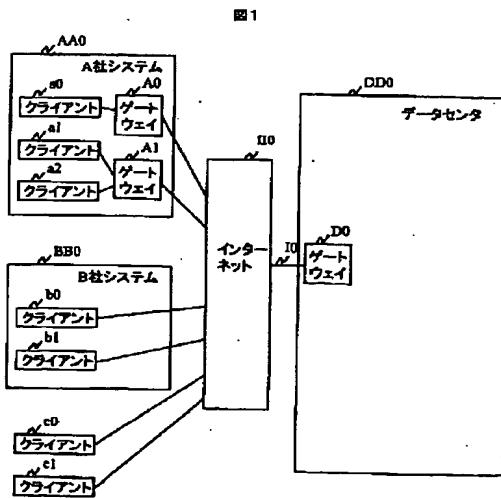
T 6 0 、 T 6 9 : ユーザ識別表、

T 7 0 : ルートファイル管理表、

2 1 0 0 、 2 4 0 0 : ユーザ条件入力画面、

2 2 0 0 、 2 3 0 0 : サービスレベル条件入力画面。

【図1】

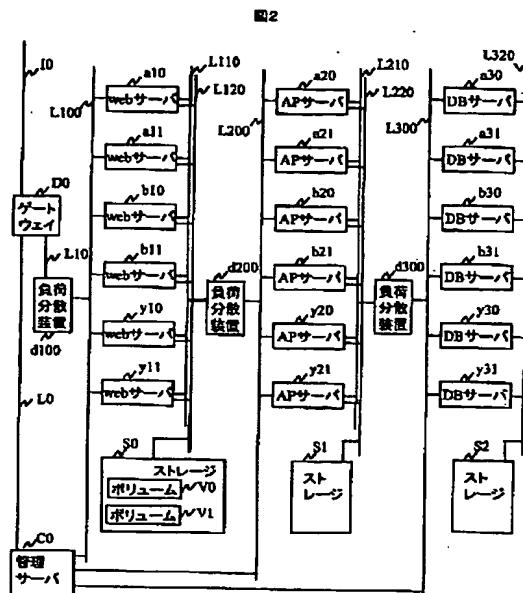


【図5】

図5

| ~ T30 | | |
|-------|--------|---------|
| UID | 仮想addr | 実addr |
| #0 | a100 | a10,a11 |
| #1 | b100 | b10,b11 |
| ~ T31 | | |
| UID | 仮想addr | 実addr |
| #0 | a100 | a10,a11 |
| | a200 | a20,a21 |
| #1 | b100 | b10,b11 |
| | b200 | b20,b21 |
| ~ T32 | | |
| UID | 仮想addr | 実addr |
| #0 | a200 | a20,a21 |
| | a300 | a30,a31 |
| #1 | b200 | b20,b21 |
| | b300 | b30,b31 |

【図2】



【図6】

図6

T33

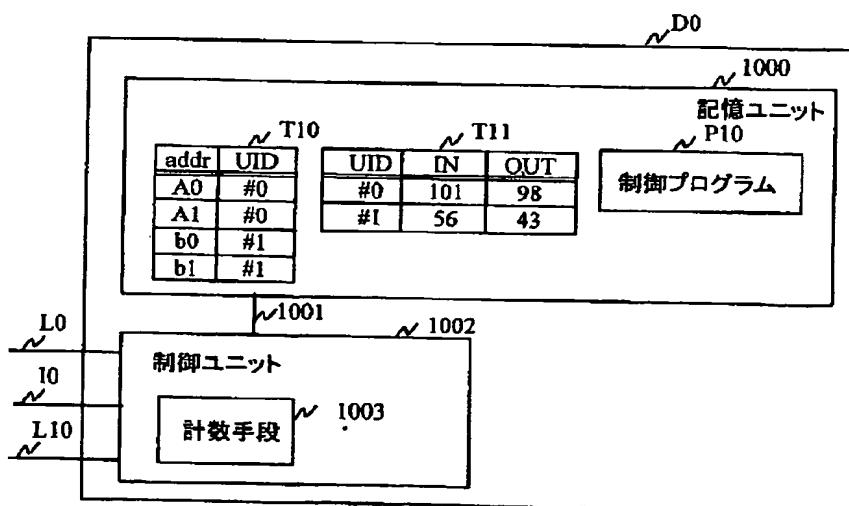
| UID | ポリューム |
|-----|-------|
| #0 | V0 |
| #1 | V1 |

【図16】

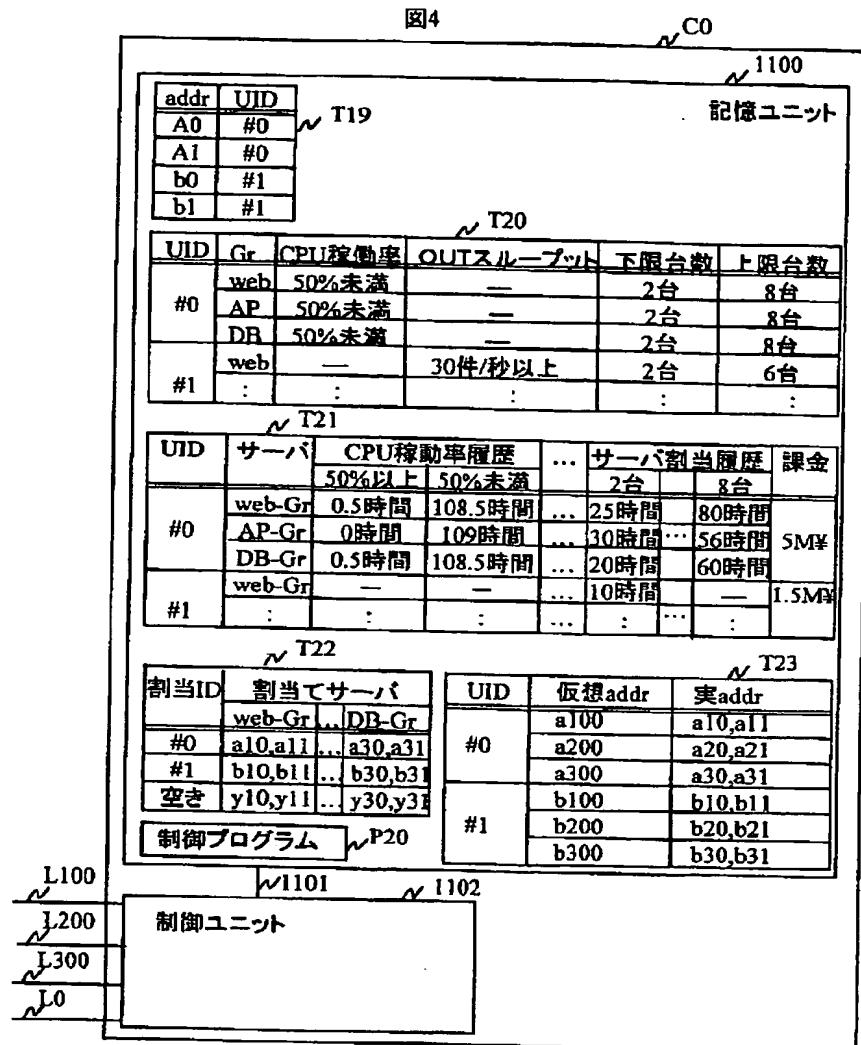
- 図16
- (1) a100' c0 data ↗1700
 - (3) a100 c0 #0 data ↗1702
 - (4) a10 c0 #0 data ↗1703
 - (13) c0 a100 #0 data ↗1712
 - (15) c0 a100 data ↗1714

【図3】

図3

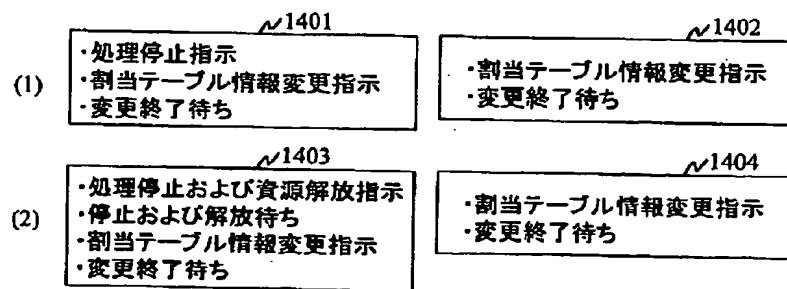


【図4】



【図9】

図9



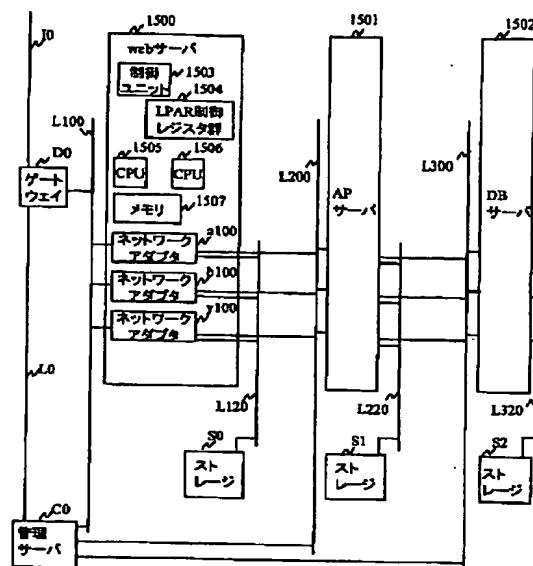
【図7】

図7

- (1) a100 a0 data ↵1200
- (2) D0 A0 data a100 a0 data ↵1201
- (3) a100 a0 #0 data ↵1202
- (4) a10 a0 #0 data ↵1203
- (5) a200 a100 #0 data ↵1204
- (6) a20 a100 #0 data ↵1205
- (7) a300 a200 #0 data ↵1206
- (8) a30 a200 #0 data ↵1207
- (9) a200 a300 #0 data ↵1208
- (10) a20 a300 #0 data ↵1209
- (11) a100 a200 #0 data ↵1210
- (12) a10 a200 #0 data ↵1211
- (13) a0 a100 #0 data ↵1212
- (14) A0 D0 data a0 a100 data ↵1213
- (15) a0 a100 data ↵1214

【図10】

図10



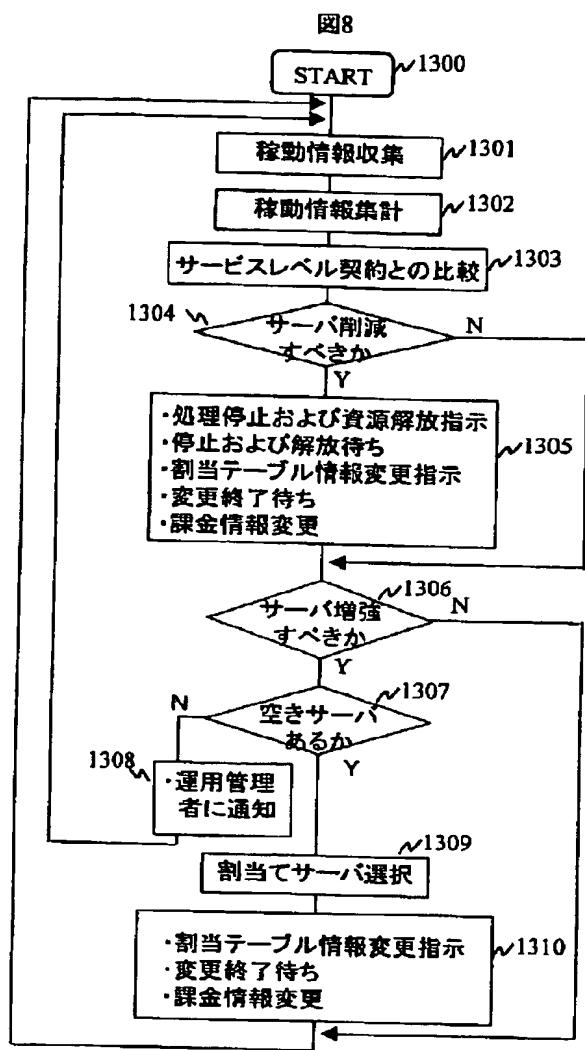
【図11】

図11

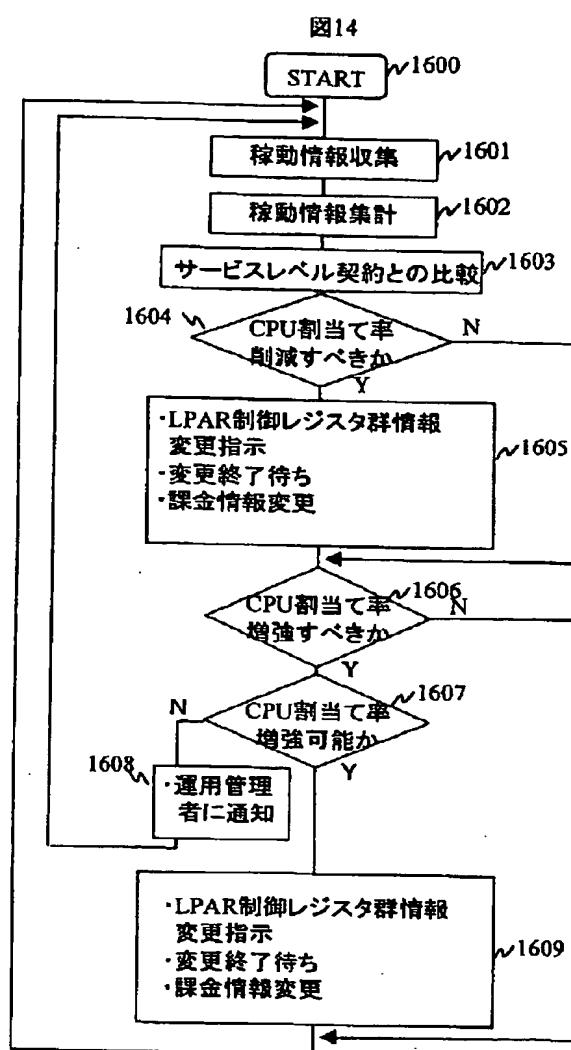
T40

| LPAR# | UID | アダプタ | メモリ割当て | | CPU割当て% |
|-------|------------|------|---------|-----|---------|
| | | | 開始オフセット | サイズ | |
| #0 | #0 | a100 | 0 | 2GB | 50% |
| #1 | #1 | b100 | 2GB | 2GB | 20% |
| #2 | not assign | y100 | 4GB | 2GB | disable |

【図8】

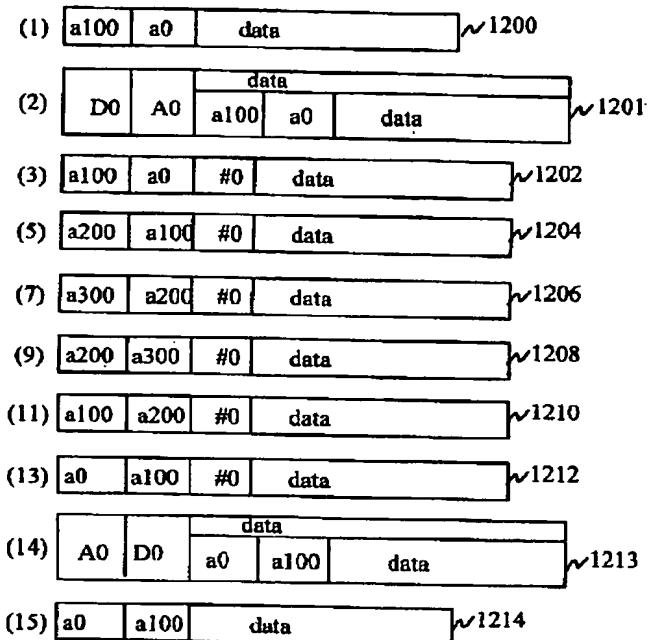


【図14】



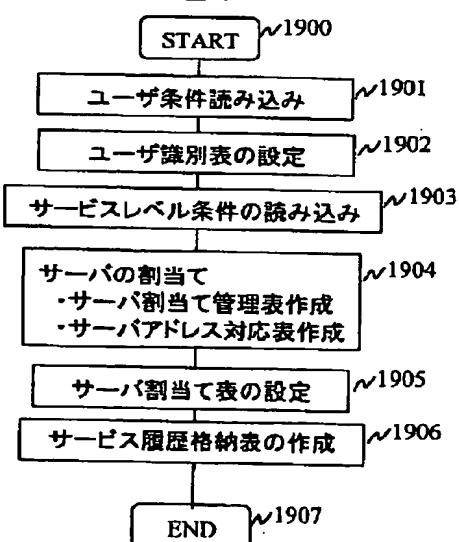
【図12】

図12



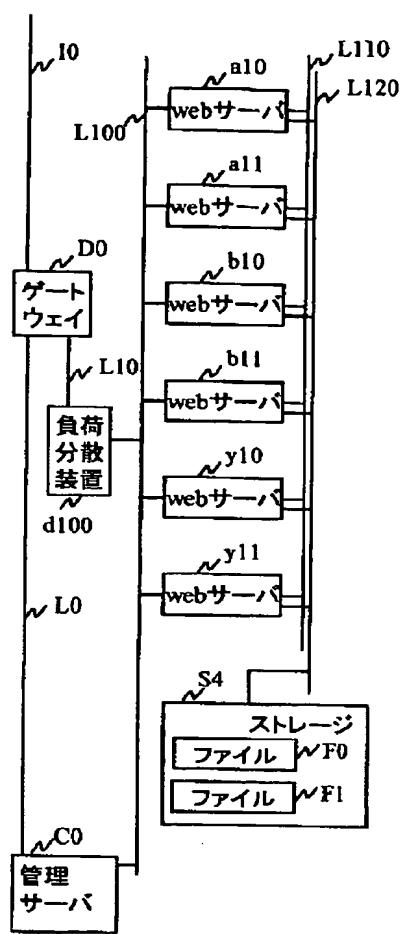
【図20】

図20



【図15】

図15



【図13】

図13

記憶ユニット T40

LPAR# UID アダプタ メモリ割当て CPU割当率 %

開始オフセット サイズ

#0 #0 a100 0 2GB 50%

#1 #1 b100 2GB 2GB 20%

#2 not assign y100 4GB 2GB disable

UID Gr CPU割当率 OUTスループット

addr UID T19 T50

#0 web 50%以上 —

A0 #0 AP 50%以上 —

A1 #0 DB 50%以上 —

b0 #1 web 20%以上 30件/秒以上

b1 #1 : : :

UID サーバ CPU使用率履歴 課金

50%未満 50%以上 T51

#0 web 0.5時間 108.5時間 ... 5M¥

AP 0時間 109時間 ...

DB 0.5時間 108.5時間 ...

#1 web — — ... 1.5M¥

AP — — ...

DB — — ...

UID サーバ 設定CPU割当率 T52

#0 web 50%

AP 50%

DB 50%

#1 web 20%

AP 20%

DB 20%

制御プログラム P20

制御ユニット

L100 L101 L102

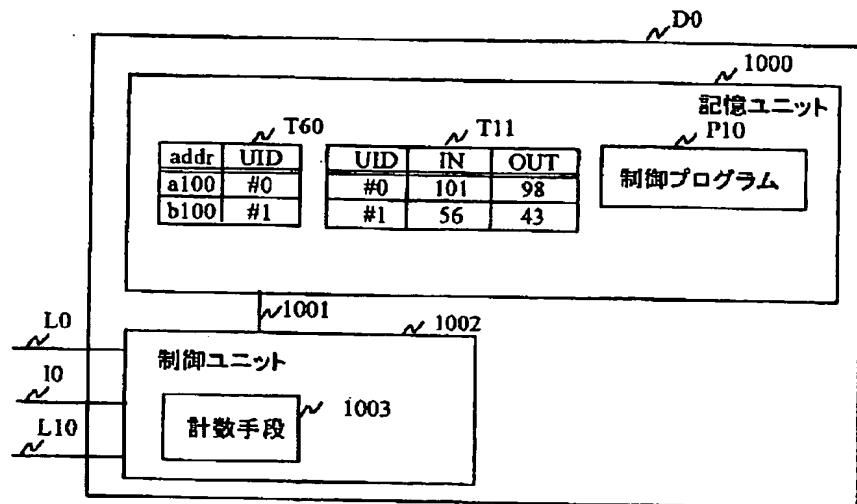
L200

L300

L0

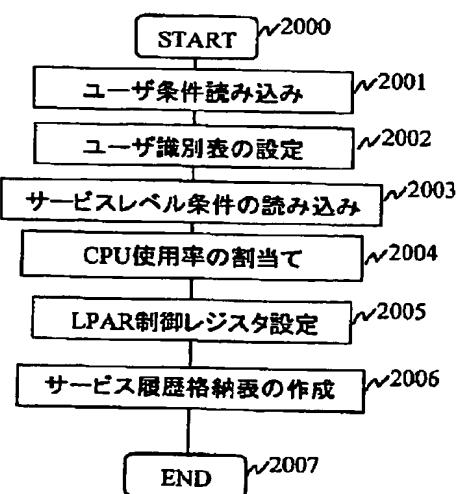
【図17】

図17

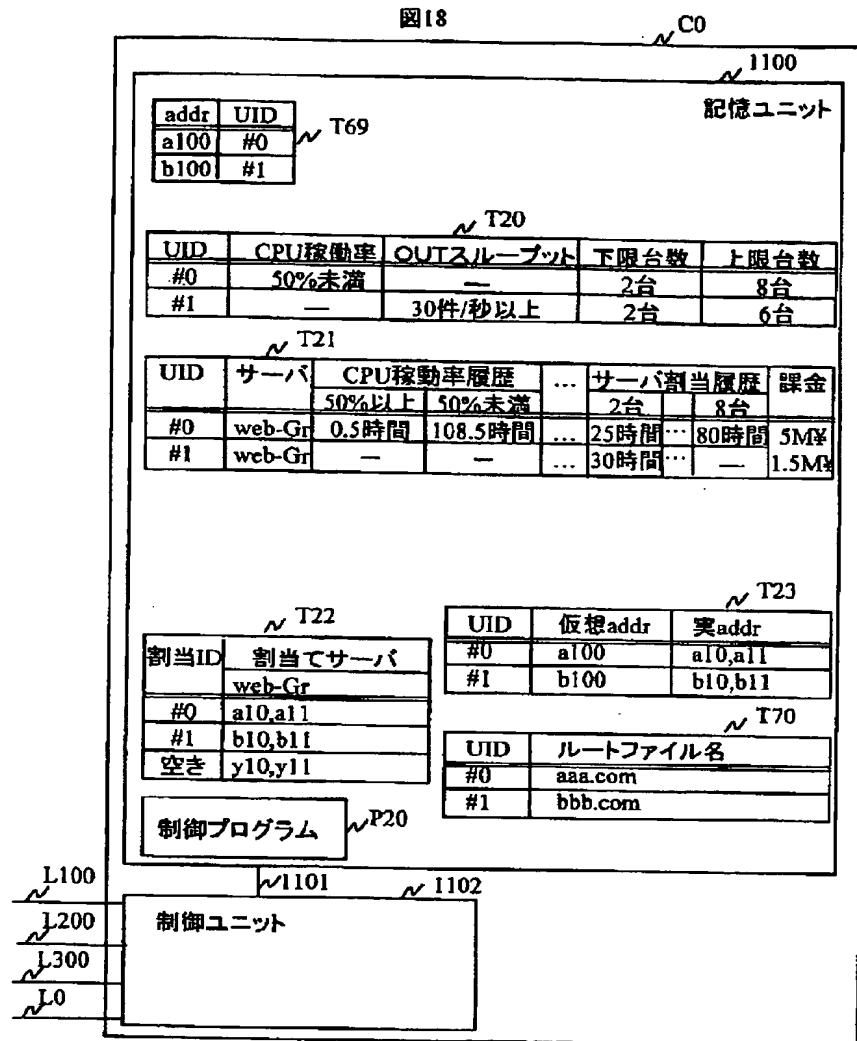


【図21】

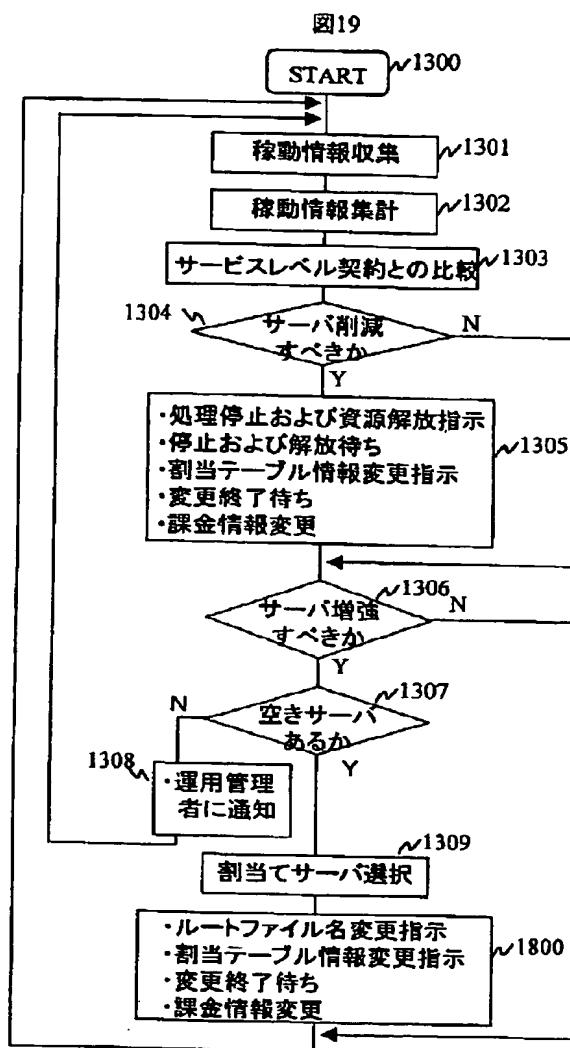
図21



【図18】



【図19】



【図22】

図22 *✓2100*

ユーザ条件入力画面

ユーザ識別条件

| | | |
|-------------|------------------------|----------------------------|
| アクセス元IPアドレス | <input type="text"/> レ | <input type="text"/> A0,A1 |
| アクセス先IPアドレス | <input type="text"/> | <input type="text"/> |
| : | <input type="text"/> | <input type="text"/> |

使用サーバ種別 使用IPアドレス

| | | |
|--------|------------------------|---------------------------|
| webサーバ | <input type="text"/> レ | <input type="text"/> a100 |
| APサーバ | <input type="text"/> レ | <input type="text"/> a200 |
| DBサーバ | <input type="text"/> レ | <input type="text"/> a300 |

【図23】

図23 *✓2200*

サービスレベル条件入力画面

使用サーバ種別 サーバ数 使用条件

| | | 下限 | 上限 | | CPU稼働率 | .. |
|--------|------------------------|------------------------|------------------------|---|----------------------------|----------------------|
| webサーバ | <input type="text"/> レ | <input type="text"/> 2 | <input type="text"/> 8 | 台 | <input type="text"/> 50%未満 | <input type="text"/> |
| APサーバ | <input type="text"/> レ | <input type="text"/> 2 | <input type="text"/> 8 | 台 | <input type="text"/> 50%未満 | <input type="text"/> |
| DBサーバ | <input type="text"/> レ | <input type="text"/> 2 | <input type="text"/> 8 | 台 | <input type="text"/> 50%未満 | <input type="text"/> |

要求性能

| | | |
|-------------------|----------------------|----------------------|
| 出力トランザクションスループット | <input type="text"/> | <input type="text"/> |
| 出力/入力スループット比 | <input type="text"/> | <input type="text"/> |
| トランザクション処理レイテンシ平均 | <input type="text"/> | <input type="text"/> |

【図24】

図24 ~ 2300

サービスレベル条件入力画面

使用サーバ種別 CPU割当て率

| | | |
|--------|---|-------|
| webサーバ | レ | 50%以上 |
| APサーバ | レ | 50%以上 |
| DBサーバ | レ | 50%以上 |

要求性能

| | |
|-------------------|----------------------|
| 出力トランザクションスループット | <input type="text"/> |
| 出力／入力スループット比 | <input type="text"/> |
| トランザクション処理レイテンシ平均 | <input type="text"/> |

【図25】

図25 ~ 2400

ユーザ条件入力画面

ユーザ識別条件

| | | |
|-------------|----------------------|----------------------|
| アクセス元IPアドレス | <input type="text"/> | <input type="text"/> |
| アクセス先IPアドレス | レ | a100 |
| : | <input type="text"/> | <input type="text"/> |

使用サーバ種別 使用IPアドレス

| | | |
|--------|----------------------|----------------------|
| webサーバ | レ | a100 |
| APサーバ | <input type="text"/> | <input type="text"/> |
| DBサーバ | <input type="text"/> | <input type="text"/> |

フロントページの続き

(72)発明者 佐川 輝俊

東京都国分寺市東恋ヶ窪一丁目280番地
株式会社日立製作所中央研究所内

(72)発明者 河辺 峻

東京都国分寺市東恋ヶ窪一丁目280番地
株式会社日立製作所中央研究所内

F ターム(参考) 5B045 GG02 GG04 JJ08

5B089 GA11 KA11 KB06 KC14 KC15

KC60 MC06

5B098 AA10 GA01 GD02 GD14

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.